6 Totale Differenzierbarkeit

Sei $U \subset \mathbb{R}$ offen. Eine Funktion $f : U \to \mathbb{R}$ ist differenzierbar in einem Punkt $x_0 \in U$ (Satz 14.6 in [EAI]) genau dann, wenn sie linear approximierbar ist in x_0 in dem Sinne, dass eine Zahl $c \in \mathbb{R}$ und eine Funktion $h : U \setminus \{x_0\} \to \mathbb{R}$ existieren mit

$$f(x) = f(x_0) + c(x - x_0) + h(x) \ (x \in U \setminus \{x_0\}) \text{ und } \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{h(x)}{x - x_0} = 0.$$

Dies ist offensichtlich äquivalent dazu, dass eine offene Umgebung $V \subset \mathbb{R}$ von $0 \in \mathbb{R}$, eine Zahl $c \in \mathbb{R}$ und eine Funktion $\varphi : V \to \mathbb{R}$ existieren mit

$$f(x_0 + \xi) = f(x_0) + c\xi + \varphi(\xi) \ (\xi \in V) \text{ und } \lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{\varphi(\xi)}{\xi} = 0.$$

Man gewinnt φ aus h, indem man $V = U - x_0$ setzt und $\varphi : V \to \mathbb{R}$ definiert durch $\varphi(\xi) = h(\xi + x_0)$ für $\xi \neq 0$ und $\varphi(0) = 0$. Die umgekehrte Implikation folgt ganz ähnlich.

Da die Abbildungen der Form

$$\mathbb{R} \to \mathbb{R}, \ \xi \mapsto c\xi \quad (c \in \mathbb{R})$$

genau die $\mathbb{R}$-linearen Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ sind, stellt die folgende Definition eine direkte Verallgemeinerung des 1-dimensionalen Differenzierbarkeitsbegriffes auf den mehrdimensionalen Fall dar.

Definition 6.1. Sei $U \subset \mathbb{R}^n$ offen und sei $x \in U$ ein Punkt in U. Eine Abbildung $f : U \to \mathbb{R}^m$ heißt *total differenzierbar* in x (oder *differenzierbar* in x), falls eine lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$, eine offene Umgebung $V \subset \mathbb{R}^n$ von $0 \in \mathbb{R}^n$ und eine Funktion $\varphi : V \to \mathbb{R}^m$ existieren mit

(i) $f(x + \xi) = f(x) + A\xi + \varphi(\xi)$ für $\xi \in V$ und

(ii) $\lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{\varphi(\xi)}{\|\xi\|} = 0$.

Man beachte, dass in der obigen Situation die Menge V immer automatisch in $U - x = \{u - x;\ u \in U\}$ enthalten ist. Indem man $\varphi(\xi) = f(x + \xi) - f(x) - A\xi$ für $\xi \in (U - x) \cap V^c$ definiert, kann man immer erreichen, dass φ auf ganz $U - x$ definiert ist.

Bemerkung 6.2. Seien in der Situation von Definition 6.1 die Funktionen $f = (f_1, \ldots, f_m)$ und $\varphi = (\varphi_1, \ldots, \varphi_m)$ gegeben durch ihre Koordinatenfunktionen. Ist $(a_{ij}) \in M(m \times n, \mathbb{R})$ die darstellende Matrix der linearen Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ bezüglich der kanonischen Basen von $\mathbb{R}^n$ und $\mathbb{R}^m$, das heißt, gilt

$$Ae_j = \sum_{i=1}^{m} a_{ij} e_i \quad (1 \leq j \leq n),$$

so ist die Gleichung (i) in Definition 6.1 äquivalent zur Gültigkeit der Gleichungen

$$f_i(x + \xi) = f_i(x) + \sum_{j=1}^{n} a_{ij}\xi_j + \varphi_i(\xi) \text{ für } \xi \in V \quad (1 \leq i \leq m).$$

Insbesondere folgt, dass eine Abbildung $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ genau dann total differenzierbar in $x \in U$ ist, wenn alle Koordinatenfunktionen $f_i : U \to \mathbb{R}$ $(1 \le i \le m)$ total differenzierbar in x sind.

Sei $V \subset \mathbb{R}^n$ eine offene Umgebung von $0 \in \mathbb{R}^n$ und $\varphi : V \to \mathbb{R}^m$ eine Funktion mit $\varphi(0) = 0$. Man schreibt

$$\varphi(\xi) = o(\|\xi\|) \quad \text{für} \quad \lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{\varphi(\xi)}{\|\xi\|} = 0$$

und schreibt abkürzend für die Gleichungen (i) und (ii) in Definition 6.1

$$f(x + \xi) = f(x) + A\xi + o(\|\xi\|).$$

Beispiel 6.3. Sei $A : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung. Dann ist A in jedem Punkt $x \in \mathbb{R}^n$ total differenzierbar, denn wegen

$$A(x + \xi) = Ax + A\xi \quad (x, \xi \in \mathbb{R}^n)$$

gelten (i) und (ii) aus Definition 6.1 mit $V = \mathbb{R}^n$ und $\varphi \equiv 0$.

Wir haben in Beispiel 5.4 (c) gesehen, dass man aus der partiellen Differenzierbarkeit einer Funktion nicht auf ihre Stetigkeit schließen kann. Als nächstes zeigen wir, dass total differenzierbare Funktionen partiell differenzierbar und stetig sind und dass die lineare Abbildung A in Gleichung (i) aus Definition 6.1 eindeutig bestimmt ist.

Satz 6.4. *Sei $U \subset \mathbb{R}^n$ offen und $x \in U$ ein Punkt in U. Es gelte*

$$f(x + \xi) = f(x) + A\xi + \varphi(\xi) \quad \textit{und} \quad \lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \frac{\varphi(\xi)}{\|\xi\|} = 0$$

wie in Definition 6.1. Sei $(a_{ij}) \in M(m \times n,\ \mathbb{R})$ die darstellende Matrix von A bezüglich der kanonischen Basen von $\mathbb{R}^n$ und $\mathbb{R}^m$. Dann gilt:

(i) f ist stetig in x,

(ii) alle Koordinatenfunktionen f_i $(i = 1, \ldots, m)$ von f sind in x partiell differenzierbar mit

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij} \quad (1 \le i \le m,\ 1 \le j \le n).$$

Insbesondere ist die lineare Abbildung A eindeutig bestimmt.

Beweis. Da die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ nach Lemma 2.20 stetig ist, folgt $\lim_{\xi \to 0} A\xi = A0 = 0$. Aus der Gültigkeit von Bedingung (ii) in Definition 6.1 folgt, dass $\lim_{\xi \to 0} \varphi(\xi) = \lim_{\substack{\xi \to 0 \\ \xi \neq 0}} \|\xi\| \frac{\varphi(\xi)}{\|\xi\|} = 0$. Damit erhält man die Stetigkeit von f in x

$$\lim_{\xi \to 0} f(x + \xi) = \lim_{\xi \to 0} (f(x) + A\xi + \varphi(\xi)) = f(x).$$

Für $i = 1, \dots, m$ und $\xi \in V$ gilt

$$f_i(x + \xi) = f_i(x) + \sum_{j=1}^{n} a_{ij}\xi_j + \varphi_i(\xi).$$

Insbesondere gilt für alle reellen Zahlen $h \neq 0$ mit genügend kleinem Absolutbetrag und für $j = 1, \dots, n$

$$\frac{f_i(x + he_j) - f_i(x)}{h} = a_{ij} + \frac{\varphi_i(he_j)}{\|he_j\|}\operatorname{sgn}(h) \overset{(h\to 0)}{\longrightarrow} a_{ij}.$$

Also ist jede Koordinatenfunktion f_i von f in jeder Koordinatenrichtung x_j partiell differenzierbar in x mit $\frac{\partial f_i}{\partial x_j}(x) = a_{ij}$. □

Da die lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ in Definition 6.1 eindeutig bestimmt ist, macht es Sinn, diese lineare Abbildung als Ableitung von f im Punkt x zu interpretieren. In der Situation der Analysis I (das heißt $n = m = 1$) bedeutet dies, dass man statt der Zahl $f'(x)$ die lineare Abbildung

$$\mathbb{R} \to \mathbb{R},\ t \mapsto f'(x)t$$

als Ableitung von f in x betrachtet.

Definition 6.5. Seien $U \subset \mathbb{R}^n$ offen, $x \in U$ ein Punkt und $f : U \to \mathbb{R}^m$ eine differenzierbare Abbildung in x mit

$$f(x + \xi) = f(x) + A\xi + \varphi(\xi) \quad \text{und} \quad \lim_{\substack{\xi\to 0\\ \xi\neq 0}} \frac{\varphi(\xi)}{\|\xi\|} = 0$$

wie in Defintion 6.1.

(a) Die eindeutig bestimmte lineare Abbildung $A : \mathbb{R}^n \to \mathbb{R}^m$ heißt das *Differential* (oder *totale Differential*) von f in x. Man schreibt

$$Df(x) \text{ (oder } f'(x)) = A$$

für das Differential von f in x.

(b) Die darstellende Matrix von A (bezüglich der kanonischen Basen von $\mathbb{R}^n$ und $\mathbb{R}^m$), das heißt die Matrix

$$\left(\frac{\partial f_i}{\partial x_j}(x)\right)_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \in M(m \times n, \mathbb{R})$$

heißt die *Jacobi-Matrix* oder *Funktionalmatrix* von f in x (geschrieben als $J_f(x)$).

Ist $A : \mathbb{R}^n \to \mathbb{R}^m$ linear, so sind die Koeffizienten der darstellenden Matrix $(a_{ij}) \in M(m \times n, \mathbb{R})$ von A die eindeutigen reellen Zahlen mit

$$Ae_j = \sum_{i=1}^{m} a_{ij}e_i \quad (1 \leq j \leq n).$$

In diesem Fall wirkt die lineare Abbildung A auf einen Vektor $x = (x_i)_{i=1}^n \in \mathbb{R}^n$ wie die Matrixmultiplikation mit der darstellenden Matrix

$$Ax = \Big(\sum_{j=1}^n a_{ij}x_j\Big)_{i=1}^m.$$

Oft identifiziert man die lineare Abbildung A mit ihrer darstellenden Matrix.

Nach Satz 6.4 impliziert die totale Differenzierbarkeit einer Funktion $f : U \to \mathbb{R}^m$ ihre partielle Differenzierbarkeit (das heißt, die partielle Differenzierbarkeit aller Koordinatenfunktionen). Die Umkehrung ist falsch, da die totale Differenzierbarkeit nach Satz 6.4 die Stetigkeit impliziert, die partielle Differenzierbarkeit aber nicht (Beispiel 5.4(c)).

Satz 6.6. *Seien $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion. Sind alle partiellen Ableitungen $D_i f$ $(i = 1, \ldots, n)$ stetig in x, so ist f total differenzierbar in x.*

Beweis. Da U offen ist, gibt es ein $\delta > 0$ mit $B_\delta(x) \subset U$. Sei $\xi \in B_\delta(0)$ und sei

$$z^{(i)} = x + \sum_{\nu=1}^i \xi_\nu e_\nu \quad (i = 0, \ldots, n).$$

Dann ist $z^{(0)} = x,\ z^{(n)} = x + \xi$ und $\|z^{(i)} - x\| \leq \|\xi\| < \delta$ für $i = 0, \ldots, n$. Der erste Mittelwertsatz der Differentialrechnung (Korollar 15.4 in [EAI]) angewendet auf die differenzierbaren Funktionen

$$g_i : [0, 1] \to \mathbb{R},\ g_i(t) = f\left(z^{(i-1)} + t\xi_i e_i\right) \quad (i = 1, \ldots, n)$$

liefert eine Zahl $\theta_i = \theta_i(\xi) \in (0, 1)$ so, dass

$$f\left(z^{(i)}\right) - f\left(z^{(i-1)}\right) = g_i(1) - g_i(0) = g_i'(\theta_i) = \xi_i D_i f\left(z^{(i-1)} + \theta_i\xi_i e_i\right).$$

Zur Berechnung der Ableitung benutze man die 1-dimensionale Kettenregel mit innerer Funktion $t \mapsto t\xi_i$. Definiert man $y^{(i)} = z^{(i-1)} + \theta_i\xi_i e_i$, so folgt

$$\begin{aligned} f(x+\xi) - f(x) = f\left(z^{(n)}\right) - f\left(z^{(0)}\right) &= \sum_{i=1}^n \left(f\left(z^{(i)}\right) - f\left(z^{(i-1)}\right)\right) \\ &= \sum_{i=1}^n D_i f\left(y^{(i)}\right)\xi_i \\ &= \sum_{i=1}^n D_i f(x)\xi_i + \varphi(\xi) \end{aligned}$$

mit $\varphi(\xi) = \sum_{i=1}^n \left(D_i f\left(y^{(i)}\right) - D_i f(x)\right)\xi_i$. Da die Funktionen $D_i f$ $(1 \leq i \leq n)$ nach Voraussetzung stetig sind in x und da

$$y^{(i)} = (x_1 + \xi_1, \ldots, x_{i-1} + \xi_{i-1},\ x_i + \theta_i\xi_i,\ x_{i+1}, \ldots, x_n) \overset{(\xi\to 0)}{\longrightarrow} x$$

konvergiert, folgt

$$\frac{|\varphi(\xi)|}{\|\xi\|} \leq \sum_{i=1}^{n} \left| D_i f(y^{(i)}) - D_i f(x) \right| \frac{|\xi_i|}{\|x\|} \overset{(\xi \to 0)}{\longrightarrow} 0.$$

Also ist f total differenzierbar in x. □

Korollar 6.7. *Sei $U \subset \mathbb{R}^n$ offen und sei $f : U \to \mathbb{R}^m$ stetig partiell differenzierbar (das heißt alle Koordinatenfunktionen $f_i : U \to \mathbb{R}$ von f seien stetig partiell differenzierbar). Dann ist f total differenzierbar und insbesondere auch stetig in jedem Punkt $x \in U$.*

Beweis. Satz 6.6 zeigt, dass alle Koordinatenfunktionen von f total differenzierbar in jedem $x \in U$ sind. Nach Bemerkung 6.2 ist f in jedem $x \in U$ total differenzierbar und nach Satz 6.4 auch stetig. □

Formal gilt für die Differentiale von total differenzierbaren Abbildungen dieselbe Kettenregel wie für die Ableitungen von Funktionen einer Veränderlichen (siehe Satz 14.10 in [EAI]).

Satz 6.8. *(Kettenregel) Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen, $x \in U$ ein Punkt und $f : U \to \mathbb{R}^m$, $g : V \to \mathbb{R}^k$ Abbildungen mit $f(U) \subset V$. Ist f differenzierbar in x und ist g differenzierbar in $f(x)$, so ist $g \circ f$ differenzierbar in x, und es gilt*

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x),$$

das heißt das totale Differential von $g \circ f$ in x ist die Komposition der totalen Differentiale von g in $f(x)$ und f in x.

Beweis. Mit $A = f'(x) : \mathbb{R}^n \to \mathbb{R}^m$ und $B = g'(f(x)) : \mathbb{R}^m \to \mathbb{R}^k$ gilt

$$\begin{aligned} f(x+u) &= f(x) + Au + \varphi(u) \text{ für } u \in U{-}x, \\ g(f(x)+v) &= g(f(x)) + Bv + \psi(v) \text{ für } v \in V{-}f(x) \end{aligned}$$

mit Funktionen $\varphi : U{-}x \to \mathbb{R}^m$, $\psi : V{-}f(x) \to \mathbb{R}^k$, für die gilt

$$\lim_{u \to 0} \frac{\varphi(u)}{\|u\|} = 0, \ \lim_{v \to 0} \frac{\psi(v)}{\|v\|} = 0.$$

Für $u \in U{-}x$ folgt

$$\begin{aligned} g \circ f(x+u) &= g(f(x+u)) = g(f(x)) + B(Au + \varphi(u)) + \psi(Au + \varphi(u)) \\ &= (g \circ f)(x) + B \circ A(u) + \chi(u) \end{aligned}$$

mit $\chi(u) = B\varphi(u) + \psi(Au + \varphi(u))$. Also genügt es zu zeigen, dass

$$\lim_{u \to 0} \frac{\chi(u)}{\|u\|} = 0.$$

Da nach Lemma 2.21

$$\left\| \frac{\chi(u)}{\|u\|} \right\| \leq \|B\| \left\| \frac{\varphi(u)}{\|u\|} \right\| + \left\| \frac{\psi(Au + \varphi(u))}{\|u\|} \right\|$$

gilt, genügt es zu zeigen, dass

$$(*) \quad \lim_{u\to 0} \frac{\psi(Au+\varphi(u))}{\|u\|} = 0$$

ist. Dazu definieren wir $\psi_0(v) = \frac{\psi(v)}{\|v\|}$ für $v \in V - f(x)$ mit $v \neq 0$ und $\psi_0(0) = 0$. Für $u \in U - x$ mit hinreichend kleiner Norm $\|u\|$, ist $\left\|\frac{\varphi(u)}{\|u\|}\right\| < 1$ und damit

$$\|\psi(Au+\varphi(u))\| \leq (\|A\|+1)\|u\| \; \|\psi_0(Au+\varphi(u))\|.$$

Indem man durch $\|u\|$ dividiert und benutzt, dass $\lim_{v\to 0}\psi_0(v) = 0$ ist, sieht man, dass Bedingung $(*)$ erfüllt ist. Dies beendet den Beweis. □

Da die darstellende Matrix einer Komposition linearer Abbildungen das Matrixprodukt der darstellenden Matrizen der linearen Abbildungen ist, gilt in der Situation von Satz 6.8

$$J_{g\circ f}(x) = J_g(f(x)) \cdot J_f(x) \quad \text{(Matrixprodukt)}.$$

Korollar 6.9. *Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen, $f = (f_1, \ldots, f_m) : U \to \mathbb{R}^m$, $g : V \to \mathbb{R}$ differenzierbare Abbildungen mit $f(U) \subset V$. Dann ist $g \circ f : U \to \mathbb{R}$ differenzierbar mit*

$$\frac{\partial(g\circ f)}{\partial x_j}(x) = \sum_{i=1}^m \frac{\partial g}{\partial x_i}(f(x))\frac{\partial f_i}{\partial x_j}(x)$$

für $x \in U$ und $j = 1, \ldots, n$ oder äquivalent

$$\text{grad } (g\circ f)(x) = (\text{grad } g(f(x)))J_f(x)$$

für alle $x \in U$.

Beweis. Nach der Bemerkung zu Satz 6.8 gilt

$$\left(\frac{\partial(g\circ f)}{\partial x_1}(x), \ldots, \frac{\partial(g\circ f)}{\partial x_n}(x)\right) = \text{grad } (g\circ f)(x) = J_{g\circ f}(x)$$

$$\begin{aligned}
&= J_g(f(x)) \cdot J_f(x) = \text{grad } g(f(x)) \cdot J_f(x) \\
&= \left(\frac{\partial g}{\partial x_1}(f(x)), \ldots, \frac{\partial g}{\partial x_m}(f(x))\right)\left(\frac{\partial f_i}{\partial x_j}(x)\right) \\
&= \left(\sum_{i=1}^m \frac{\partial g}{\partial x_i}(f(x))\frac{\partial f_i}{\partial x_1}(x), \ldots, \sum_{i=1}^m \frac{\partial g}{\partial x_i}(f(x))\frac{\partial f_i}{\partial x_n}(x)\right).
\end{aligned}$$ □

Korollar 6.10. *Seien $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ offen und $k \in \mathbb{N}$. Ist $f \in C^k(U)$ mit $f(U) \subset V$ und ist $g \in C^k(V)$, so ist $g \circ f \in C^k(V)$.*

Beweis. Mit der Produktregel und Induktion nach k folgt, dass für zwei Funktionen $u, v \in C^k(U)$ das Produkt $uv \in C^k(U)$ gehört und dass die partiellen Ableitungen von uv der Ordnung $\leq k$ endliche Summen von Produkten aus einer partiellen Ableitung von u und einer parteillen Ableitung von v jeweils der

Ordnung $\leq k$ sind. Damit folgt die Behauptung von Korollar 6.10 durch Induktion nach k. Der Induktionsanfang $k = 1$ folgt direkt aus Korollar 6.9. Sei $k \geq 2$ und die Behauptung gezeigt für $k-1$. Für f, g wie in Korollar 6.10 und $i_1, \ldots, i_k \in \{1, \ldots, n\}$ ist dann nach Korollar 6.9, der Induktionsvoraussetzung und der obigen Bemerkung über Produkte

$$D_{i_1}(g \circ f) = \sum_{i=1}^{m} (\partial_i g) \circ f \cdot (\partial_{i_1} f) \in C^{k-1}(U).$$

Also existiert $D_{i_k} \cdots D_{i_2}(D_{i_1}(g \circ f))$ und ist stetig. □

Die partiellen Ableitungen sind Ableitungen in Richtung der Koordinatenachsen. Entsprechend kann man Ableitungen in Richtung beliebiger Vektoren im $\mathbb{R}^n$ bilden.

Definition 6.11. (Richtungsableitungen) Seien $U \subset \mathbb{R}^n$ offen, $x \in U,\ v \in \mathbb{R}^n$ mit $\|v\| = 1$ und $f : U \to \mathbb{R}$ eine Funktion. Man nennt (falls dieser Limes existiert)

$$D_v f(x) = \frac{d}{dt} f(x + tv)|_{t=0} = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t} \in \mathbb{R}$$

die *Richtungsableitung* von f im Punkt x in Richtung v.

Für $v = e_i$ $(1 \leq i \leq n)$ ist offensichtlich $D_{e_i} f(x) = D_i f(x)$.

Satz 6.12. *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ differenzierbar. Für $x \in U$ und $v \in \mathbb{R}^n$ mit $\|v\| = 1$ ist*

$$D_v f(x) = \langle \text{grad } f(x), v \rangle = f'(x)(v).$$

Beweis. Nach der Kettenregel (Satz 6.9) ist die Funktion

$$W = \{t \in \mathbb{R};\ x + tv \in U\} \to \mathbb{R},\ t \mapsto f(x + tv)$$

differenzierbar mit

$$\frac{d}{dt} f(x+tv) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+tv) \frac{d(x_i + tv_i)}{dt} = \langle \text{grad } f(x+tv), v \rangle = \sum_{i=1}^{n} v_i f'(x)(e_i) = f'(x)(v)$$

für alle $t \in W$. □

Bemerkung 6.13. Ist grad $f(x) \neq 0$ in Satz 6.12, so gilt

$$D_v f(x) = \langle \text{grad } f(x), v \rangle = \cos\theta \|\text{grad } f(x)\|,$$

wobei $\theta \in [0, \pi]$ der Schnittwinkel zwischen v und grad $f(x)$ ist. Die Richtungsableitung in x wird maximal in Richtung $v = \frac{\text{grad } f(x)}{\|\text{grad } f(x)\|}$.

Sei $I \subset \mathbb{R}$ ein Intervall positiver Länge und $f : I \to \mathbb{R}$ stetig differenzierbar. Sind $x, \xi \in \mathbb{R}$ mit $x, x+\xi \in I$, so folgt aus der Substitutionsregel mit der inneren Funktion $\varphi(t) = x + t\xi$ $(t \in [0,1])$

$$f(x+\xi) - f(x) = \int_{x=\varphi(0)}^{x+\xi=\varphi(1)} f'(u)du = \int_0^1 f'(\varphi(t))\varphi'(t)dt = \xi \int_0^1 f'(x+t\xi)dt.$$

Diese Version des Mittelwertsatzes der Differentialrechnung gilt auch mehrdimensional. Dabei ersetzen wir die Ableitung von f unter dem letzten Integral durch die Jacobi-Matrix von f. Die dabei auftretenden Integrale matrixwertiger stetiger Funktionen definiert man koeffizientenweise.

Definition 6.14. Seien $m, n \in \mathbb{N}^*$, $a, b \in \mathbb{R}$ mit $a < b$ und sei

$$A : [a,b] \to M(m \times n, \mathbb{R}),\ t \mapsto A(t) = (a_{ij}(t))$$

stetig. Dann definiert man

$$\int_a^b A(t)dt = \left(\int_a^b a_{ij}(t)dt \right) \in M(m \times n, \mathbb{R}).$$

Man beachte dabei, dass nach Bemerkung 2.22 alle Koeffizientenfunktionen $[a,b] \to \mathbb{R},\ t \mapsto a_{ij}(t)$ stetig sind.

Satz 6.15. *(Mittelwertsatz) Seien $U \subset \mathbb{R}^n$ offen, $f \in C^1(U, \mathbb{R}^m)$, $x \in U$ und $\xi \in \mathbb{R}^n$ so, dass $x + t\xi \in U$ für alle $t \in [0,1]$. Dann gilt*

$$f(x+\xi) - f(x) = \left(\int_0^1 J_f(x+t\xi) \right) \cdot \xi.$$

Beweis. Sei $f : U \to \mathbb{R}^m$, $f(x) = (f_1(x), \ldots, f_m(x))$ stetig differenzierbar. Nach der Kettenregel (Korollar 6.9) sind die Funktionen

$$g_i : [0,1] \to \mathbb{R},\ t \mapsto f_i(x+t\xi) \quad (1 \le i \le m)$$

stetig differenzierbar. Mit dem Hauptsatz der Differential- und Integralrechnung (Satz 17.7 in [EAI]) und der Kettenregel folgt, dass

$$\begin{aligned} f_i(x+\xi) - f_i(x) = g_i(1) - g_i(0) &= \int_0^1 g_i'(t)dt \\ &= \int_0^1 \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x+t\xi) \frac{d}{dt}(x_j + t\xi_j)dt \\ &= \sum_{j=1}^n \left(\int_0^1 \frac{\partial f_i}{\partial x_j}(x+t\xi) \right) dt \cdot \xi_j \end{aligned}$$

gleich der i-ten Zeile des Produktes

$$\left(\int_0^1 \frac{\partial f_\mu}{\partial x_\nu}(x+t\xi)dt \right)_{\substack{1 \le \mu \le m \\ 1 \le \nu \le n}} (\xi_\nu)_{\nu=1}^n = \left(\int_0^1 J_f(x+t\xi)dt \right) \cdot \xi$$

ist. $\square$

Für Anwendungen des Mittelwertsatzes ist es nützlich, eine Verallgemeinerung der Standardabschätzung für Riemann-Integrale (Satz 16.13(b) in [EAI]) zur Verfügung zu haben.

Lemma 6.16. *Für stetige Funktionen $f : [a, b] \to \mathbb{R}^n,\ A : [a, b] \to M(m \times n, \mathbb{R})$ gilt*

$$\left\| \int_a^b f(t)dt \right\| \leq \int_a^b \|f(t)\| \, dt$$

und

$$\left\| \int_a^b A(t)dt \right\| \leq \int_a^b \|A(t)\| \, dt,$$

wobei auch das erste Integral komponentenweise definiert ist.

Beweis. Für eine Teilung $T = (t_i)_{i=0}^r$ von $[a, b]$ und eine Zwischenfolge $Z = (z_i)_{i=1}^r$ von T definieren wir wie in der Analysis I (Definition 16.10 in [EAI]) die zugehörige Riemann-Summe der stetigen Funktion $f : [a, b] \to \mathbb{R}^n$ durch

$$S(f, T, Z) = \sum_{i=1}^r (t_1 - t_{i-1}) f(z_i) \in \mathbb{R}^n.$$

Dann ist $S(f, T, Z) = (S(f_j, T, Z))_{j=1}^n$. Sei $(T_k)_{k\geq1}$ eine Folge von Teilungen des Intervalls $[a, b]$ mit $\lim_{k\to\infty} \omega(T_k) = 0$ und $(Z_k)_{k\geq1}$ eine Folge von Zwischenfolgen Z_k von T_k. Nach Satz 16.11 in [EAI] und Lemma 2.2 gilt

$$\int_a^b f \ dt = \left(\int_a^b f_j dt \right)_{j=1}^n = \lim_{k\to\infty} (S(f_j, T_k, Z_k))_{j=1}^n = \lim_{k\to\infty} S(f, T_k, Z_k).$$

Da die Normfunktion $\| \cdot \|$ stetig ist (Beispiel 2.13 (b)), folgt

$$\left\| \int_a^b f \ dt \right\| = \lim_{k\to\infty} \|S(f, T_k, Z_k)\| \leq \lim_{k\to\infty} S(\|f\|, T_k, Z_k) = \int_a^b \|f\| dt,$$

wobei wir die Dreiecksungleichung für die euklidische Norm $\| \cdot \|$ und noch einmal Satz 16.11 aus [EAI] (diesesmal für die stetige Funktion $\|f\|$) benutzt haben.
Da auch Konvergenz in $M(m\times n, \mathbb{R})$ äquivalent zur koeffizientenweisen Konvergenz ist (Bemerkung 2.22), kann man die Integralabschätzung im matrixwertigen Fall ganz genauso beweisen. □

Kombiniert man den Mittelwertsatz mit der obigen Integralabschätzung, so hält man die *Mittelwertabschätzung* der mehrdimensionalen Differentialrechnung.

Korollar 6.17. *(Mittelwertabschätzung)*
Seien $U \subset \mathbb{R}^n$ offen, $f \in C^1(U, \mathbb{R}^m),\ x \in U,\ \xi \in \mathbb{R}^n$ mit $x + t\xi \in U$ für alle $t \in [0, 1]$. Dann gilt

$$\|f(x + \xi) - f(x)\| \leq M\|\xi\|$$

mit

$$M = \sup_{t\in[0,1]} \|J_f(x + t\xi)\| = \sup_{t\in[0,1]} \|f'(x + t\xi)\| < \infty.$$

Beweis. Mit dem Mittelwertsatz (Satz 6.15) sowie Lemma 2.21 und Lemma 6.16 enthält man

$$\|f(x+\xi) - f(x)\| = \left\| \left(\int_0^1 J_f(x+t\xi)dt \right) \cdot \xi \right\| \leq \left\| \int_0^1 J_f(x+t\xi)dt \right\| \ \|\xi\|$$

$$\leq \int_0^1 \|J_f(x+t\xi)\|dt \ \|\xi\| \leq M\|\xi\|.$$

Da $f \in C^1(U, \mathbb{R}^m)$ ist, ist die Funktion $[0,1] \to M(m \times n, \mathbb{R}),\ t \mapsto J_f(x+t\xi)$ nach Bemerkung 2.22 stetig. Die Stetigkeit bleibt erhalten, wenn man die Matrixnorm (Lemma 2.21) hinter diese Funktion schaltet. Also ist $M < \infty$. Da die Norm einer Matrix $A \in M(m \times n, \mathbb{R})$ definiert wurde als die Operatornorm des zugehörgien Multiplikationsoperators $\mathbb{R}^n \to \mathbb{R}^m,\ x \mapsto Ax$ (Lemma 2.21), folgt die Gleichheit der beiden Suprema. □

Literatur

[EAI] Eschmeier, J., Analysis I, Vorlesungsskript, Universität des Saarlandes, 2013.